

Exploiting Equitable Partitions for Efficient Block Triangularization

Mario Thüne

Abstract

In graph theory a partition of the vertex set of a graph is called equitable if for all pairs of cells all vertices in one cell have an equal number of neighbours in the other cell. Considering the implications for the adjacency matrix one may generalize that concept as a block partition of a complex square matrix s.t. each block has constant row sum. It is well known that replacing each block by its row sum yields a smaller matrix whose multiset of eigenvalues is contained in the initial spectrum. We generalize this approach to weighted row sums and rectangular matrices and derive an efficient unitary transformation which approximately block triangularizes a matrix w.r.t. an arbitrary partition. Singular values and Hermiticity (if present) are preserved. The approximation is exact in the equitable case and the error can be bounded in terms of unitarily invariant matrix norms.

1 Introduction

1.1 Equitable Partitions

Let Γ be a (multi-)graph and let \mathbf{A} be its adjacency matrix, whose entries a_{vw} are the number of edges connecting vertices v and w . Let $\Pi = (c_1, \dots, c_k)$ be a partition of the vertex set of Γ into k cells, inducing a block partition of \mathbf{A} , i.e. a simultaneous (disjoint and exhaustive) partition of its rows and columns. It is convenient to define an *indicator matrix* of a partition as

Definition 1.

$$\mathbf{B} = (b_{vi}) \in \{0, 1\}^{N \times k} \text{ with } b_{vi} = \begin{cases} 1 & , \text{ if item } v \text{ is in cell } i \\ 0 & , \text{ else.} \end{cases}$$

The partition Π is called equitable if all vertices of Γ in the same cell have the same number of neighbours in any cell. Equivalently, we may call it equitable if each induced submatrix of \mathbf{A} has constant row sum. The equitable partitions of \mathbf{A} ordered by refinement form a lattice which contains the trivial equitable partition, in which every cell has size exactly one, as the minimum.

From the definition it follows that a partition of \mathbf{A} is equitable if and only if there exists a matrix $\Theta = (\theta_{ij})$ s.t.

$$\mathbf{A}\mathbf{B} = \mathbf{B}\Theta \quad \text{i.e.} \quad \forall i, j \in \{1, \dots, k\} \quad \forall u \in \{1, \dots, n_i\} \quad \sum_v^{n_j} \mathbf{A}_{ij,uv} = \theta_{ij}. \quad (1)$$

The matrix Θ is called the quotient of the partition. Its entries θ_{ij} are the constant row sums of the matrix blocks A_{ij} induced by the cells c_i and c_j , which are the number of edges connecting a fixed vertex in c_i to vertices in c_j .

1.2 Applications in Graph and Matrix Theory

The notion of equitable partitions was developed in graph theory. In network analysis the same concept is also known as *exact coloration* [3] or *exact role assignment* [20]. It is closely related to *graph fibration* [2] and arises naturally in the context of graph automorphisms problems since every non trivial automorphism induces a non trivial equitable partition. As graph invariants which can be searched for quickly using quite efficient algorithms, equitable partitions are useful in attacking graph isomorphism problems. In that context they are also known as *1-dimensional Weisfeiler-Lehman stabilizers* [5].

Block partitioned matrices s.t. each block has constant row sum are called *block-stochastic matrices*. In the context of markov chains the technique of *lumping* exploits equitable partitions in order to reduce the number of states [4]. The quotient Θ is also known as the *front divisor* [7]. Famously, the spectrum of Θ , called the main spectrum, is a subset of the spectrum of \mathbf{A} since the columns of \mathbf{B} span an invariant subspace if (1) holds. Therefore, there is a similarity transformation which 2×2 block triangularizes \mathbf{A} s.t. one diagonal block is the quotient. Such a transformation can be constructed and applied efficiently in a way provided in [15], [6]. The block triangularization method given below differs from that approach in order to fit in a generalized framework of equitability and provides efficient unitary transformations.

1.3 Aim and Outline

We will generalize the notion of ordinary equitable partitions to arbitrary weighted partitions of the rows and columns of complex matrices. According to a given partition we derive an efficient and stable unitary similarity transformation in order to 2×2 block triangularize the matrix up to an error term, which is minimized w.r.t. to several matrix norms and vanishes if and only if exact equitability holds. The transformation can be computed in $O(N)$ and applied in $O(N^2)$. It can be further generalized enabling the application to rectangular matrices while maintaining the unitarity property. However the further generalized transformation does only preserve the singular values, but (in general) not the spectrum.

Despite offering insights into the structure of objects represented by a graph

or a matrix our notion of equitability and its corresponding transformation may be used for compression and for preprocessing eigen and singular value problems. Although describing our transformation as an efficient compression method may seem to suggest that the exploited structure is, in a sense, rare, the concept of equitable partitions, as indicated above, is rather common in various applications, where it is found directly in the studied problem or as an interesting exceptional or ideal case. The usefulness might be increased in particular by the fact that deviations from an exact equitability may be allowed within our framework.

In order to get used to the concept and some notation, we briefly discuss in section (2) the special case of an ordinary unweighted equitable partition of a complex square matrix including the derivation of the associated efficient unitary block triangularization and we give an example. In the main part, section (3), we consider weighted not necessarily equitable partitions introducing the deviation matrix and give our main theorem. In section (4) we consider non exact equitability as an eigenvalue perturbation, give a short overview of several other known generalizations of equation (1), and briefly consider the problem of finding an equitable partition. Our further generalized version of the concept applicable to rectangular matrices can be found in the appendix. Throughout the article we use the apostrophe to denote the complex conjugated transpose without distinguishing between real and complex operands and we utilize the following notation

Definition 2. Let $n \in \mathbb{N}$. $\mathbf{j}_n = (\underbrace{1, \dots, 1}_{n \text{ times}})'$ and $\mathbf{f}_n = (1, \underbrace{0, \dots, 0}_{(n-1) \text{ times}})'$.

2 Unweighted Equitable Partitions

2.1 Indicator Matrix and Quotient

Let $\mathbf{A} \in \mathbb{C}^{N \times N}$ and let $\Pi = (c_1, \dots, c_k)$ be a simultaneous (disjoint and exhaustive) partition of its rows and columns with *indicator matrix* \mathbf{B} as in definition (1). Let $\mathbf{A}_{ij} \in \mathbb{C}^{n_i \times n_j}$ be the matrix block in \mathbf{A} induced by row cell c_i and column cell c_j . Let n_i be the size of the cell c_i and let

$$\mathbf{N} = (\mathbf{B}'\mathbf{B})^{\frac{1}{2}} = \text{diag}(\sqrt{n_1}, \dots, \sqrt{n_k}) \quad (2)$$

We introduce the front quotient, the rear quotient and the Rayleigh quotient respectively as

$$\mathbf{E}^- = \mathbf{N}^{-2}\mathbf{B}'\mathbf{A}\mathbf{B}, \quad \mathbf{E}^+ = \mathbf{B}'\mathbf{A}\mathbf{B}\mathbf{N}^{-2}, \quad \mathbf{E}^0 = \mathbf{N}^{-1}\mathbf{B}'\mathbf{A}\mathbf{B}\mathbf{N}^{-1} \quad (3)$$

We call \mathbf{A} *front equitable* (i) and respectively *rear equitable* (ii) w.r.t. \mathbf{B} if

$$(i) \quad \mathbf{A}\mathbf{B} = \mathbf{B}\mathbf{E}^- \quad , \quad (ii) \quad \mathbf{B}'\mathbf{A} = \mathbf{E}^+\mathbf{B}'. \quad (4)$$

It is easy to see that for Hermitian matrices row equitability and column equitability imply each other. For the rest of this section we assume front equitability, i.e.

$$\mathbf{A}_{ij}\mathbf{j}_{n_j} = e_{ij}^-\mathbf{j}_{n_i} \quad \forall i, j \in \{1, \dots, k\}. \quad (5)$$

2.2 Block Triangularization

In order to block triangularize \mathbf{A} according to \mathbf{B} we utilize the Householder matrices

$$\mathbf{H}_i = \mathbf{I}_{n_i} - 2 \frac{\mathbf{y}_i \mathbf{y}_i'}{\mathbf{y}_i' \mathbf{y}_i}, \quad \mathbf{y}_i = \mathbf{j}_i + \sqrt{n_i} \mathbf{f}_{n_i}. \quad (6)$$

The following useful relations are easily verified

$$\mathbf{H}_i \mathbf{f}_{n_i} = -\frac{1}{\sqrt{n_i}} \mathbf{j}_{n_i}, \quad \mathbf{H}_i' \mathbf{j}_{n_i} = -\sqrt{n_i} \mathbf{f}_{n_i}. \quad (7)$$

In order to simplify notations but w.l.o.g. we assume *suitable indexing* which means that \mathbf{A} and \mathbf{B} are indexed in such a way that for u in cell c_i and v in cell c_j it holds that $i < j$ implies $u < v$. Then our proposed transformation of \mathbf{A} can be written conveniently in matrix form using the matrix

$$\tilde{\mathbf{H}} = \text{diag}(\mathbf{H}_1, \dots, \mathbf{H}_k), \quad (8)$$

which is explicitly block diagonal and, according to (6), unitary.

$$\tilde{\mathbf{A}} = \tilde{\mathbf{H}}' \mathbf{A} \tilde{\mathbf{H}} = \begin{pmatrix} \tilde{\mathbf{A}}_{11} & \cdots & \tilde{\mathbf{A}}_{1k} \\ \vdots & \ddots & \vdots \\ \tilde{\mathbf{A}}_{k1} & \cdots & \tilde{\mathbf{A}}_{kk} \end{pmatrix} \quad \text{with} \quad \tilde{\mathbf{A}}_{ij} = \mathbf{H}_i' \mathbf{A}_{ij} \mathbf{H}_j. \quad (9)$$

By (7) and (5) there exists a matrix $\mathbf{E} = (e_{ij})$ s.t.

$$\tilde{\mathbf{A}}_{ij} \mathbf{f}_j = e_{ij} \mathbf{f}_{n_i}, \quad (10)$$

which immediately shows that each $\tilde{\mathbf{A}}_{ij}$ is block triangular with the left upper block being the scalar e_{ij} . Therefore, there is a readily available, in general not unique permutation matrix $\mathbf{\Omega}$ such that

$$\hat{\mathbf{A}} = \mathbf{\Omega}' \tilde{\mathbf{A}} \mathbf{\Omega} = \begin{pmatrix} \mathbf{E} & \mathbf{D} \\ \mathbf{0} & \mathbf{F} \end{pmatrix} \quad (11)$$

is explicitly block triangular. Since the applied transformations are unitary, the spectrum and the singular values of \mathbf{A} are preserved. We will refer to \mathbf{F} , which in general depends on the indexing of \mathbf{A} and on $\mathbf{\Omega}$, as a *factor*. One shows that all factors are unitarily equivalent and that by similarity

$$\sigma(\mathbf{A}) = \sigma(\mathbf{E}) + \sigma(\mathbf{F}). \quad (12)$$

Additionally, if \mathbf{v} is an eigenvector of $\hat{\mathbf{A}}$ then $\tilde{\mathbf{H}}\mathbf{\Omega}\mathbf{v}$ is an eigenvector of \mathbf{A} to the same eigenvalue. One also shows that \mathbf{D} vanishes if and only if rear equitability holds. The computational costs for the transformation $\tilde{\mathbf{H}}$ are of order $O(n_i n_j)$ on each subblock for we apply only matrix vector multiplication and matrix addition since \mathbf{H}_i is a rank one update of the identity. Therefore, the total costs are of order $O(N^2)$. Since $\tilde{\mathbf{H}}\mathbf{\Omega}$ is unitary, Hermiticity (if present) of \mathbf{A} is preserved. Numeric stability is supported by using Householder matrices. Note that in this section we constructed $\tilde{\mathbf{H}}$ s.t. $\mathbf{E} = \mathbf{E}^0$. In the general case those two matrices are unitarily equivalent but not necessarily identical.

2.3 Example

Let

$$\mathbf{A}_0 = \begin{pmatrix} 1 & 2 & 3 & 3 & 3 & 2 \\ 2 & 4 & 3 & 1 & 2 & 1 \\ 3 & 3 & 1 & 4 & 1 & 1 \\ 3 & 1 & 4 & 0 & 2 & 3 \\ 3 & 2 & 1 & 2 & 3 & 2 \\ 2 & 1 & 1 & 3 & 2 & 4 \end{pmatrix} \quad \text{and} \quad \mathbf{P}_0 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}$$

One verifies that \mathbf{A}_0 is (unweighted) front equitable w.r.t. $\Pi_0 = (1|2, 6|3, 4, 5)$. Using the permutation \mathbf{P}_0 we can transform it into the suitably indexed form

$$\mathbf{A} = \mathbf{P}_0' \mathbf{A}_0 \mathbf{P}_0 = \begin{pmatrix} 1 & 2 & 2 & 3 & 3 & 3 \\ 2 & 4 & 1 & 1 & 2 & 3 \\ 2 & 1 & 4 & 3 & 2 & 1 \\ 3 & 1 & 3 & 0 & 2 & 4 \\ 3 & 2 & 2 & 2 & 3 & 1 \\ 3 & 3 & 1 & 4 & 1 & 1 \end{pmatrix},$$

which is (unweighted) front equitable w.r.t. $\Pi = (1|2, 3|4, 5, 6)$ with front quotient

$$\mathbf{E}^- = \begin{pmatrix} 1 & 4 & 9 \\ 2 & 5 & 6 \\ 3 & 4 & 6 \end{pmatrix}.$$

One may employ

$$\begin{aligned} \mathbf{H}_1 &= \mathbf{H}(\mathbf{j}_1) = -1, \\ \mathbf{H}_2 &= \mathbf{H}(\mathbf{j}_2) = -\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \\ \mathbf{H}_3 &= \mathbf{H}(\mathbf{j}_3) = -\frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 1 & 1 \\ 1 & \frac{1+\sqrt{3}}{-2} & \frac{1-\sqrt{3}}{-2} \\ 1 & \frac{1-\sqrt{3}}{-2} & \frac{1+\sqrt{3}}{-2} \end{pmatrix} \end{aligned}$$

and $\tilde{\mathbf{H}} = \text{diag}(\mathbf{H}_1, \mathbf{H}_2, \mathbf{H}_3)$ to transform \mathbf{A} s.t.

$$\tilde{\mathbf{A}} = \tilde{\mathbf{H}}' \mathbf{A} \tilde{\mathbf{H}} = \begin{pmatrix} 1 & \frac{4}{\sqrt{2}} & 0 & \frac{9}{\sqrt{3}} & 0 & 0 \\ \frac{4}{\sqrt{2}} & 5 & 0 & 6\frac{\sqrt{2}}{\sqrt{3}} & 0 & 0 \\ 0 & 0 & 3 & 0 & -3+\sqrt{3} & -3-\sqrt{3} \\ \frac{9}{\sqrt{3}} & 6\frac{\sqrt{2}}{\sqrt{3}} & 0 & 6 & 0 & 0 \\ 0 & 0 & -3+\sqrt{3} & 0 & \sqrt{3}-1 & -6 \\ 0 & 0 & -3-\sqrt{3} & 0 & -6 & -\sqrt{3}-1 \end{pmatrix}.$$

Using the permutation $\mathbf{\Omega} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$ we obtain the matrix

$$\hat{\mathbf{A}} = \mathbf{\Omega}' \tilde{\mathbf{A}} \mathbf{\Omega} = \begin{pmatrix} 1 & \frac{4}{\sqrt{2}} & \frac{9}{\sqrt{3}} & 0 & 0 & 0 \\ \frac{4}{\sqrt{2}} & 5 & 6\frac{\sqrt{2}}{\sqrt{3}} & 0 & 0 & 0 \\ \frac{9}{\sqrt{3}} & 6\frac{\sqrt{2}}{\sqrt{3}} & 6 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & -3+\sqrt{3} & -3-\sqrt{3} \\ 0 & 0 & 0 & -3+\sqrt{3} & \sqrt{3}-1 & -6 \\ 0 & 0 & 0 & -3-\sqrt{3} & -6 & -\sqrt{3}-1 \end{pmatrix},$$

which is explicitly reducible. Since \mathbf{A} is Hermitian, the (unweighted) partition Π induces front and rear equitability and we actually obtain a block diagonal form. Note that both blocks are Hermitian but the front quotient \mathbf{E}^- is not. One verifies that

$$\mathbf{E} = \begin{pmatrix} 1 & \frac{4}{\sqrt{2}} & \frac{9}{\sqrt{3}} \\ \frac{4}{\sqrt{2}} & 5 & 6\frac{\sqrt{2}}{\sqrt{3}} \\ \frac{9}{\sqrt{3}} & 6\frac{\sqrt{2}}{\sqrt{3}} & 6 \end{pmatrix} = \text{diag}(1, 2, 3)^{\frac{1}{2}} \mathbf{E}^- \text{diag}(1, 2, 3)^{-\frac{1}{2}}.$$

Let \mathbf{F} denote the lower diagonal block of $\hat{\mathbf{A}}$. Let $\mathbf{V}_\mathbf{E}$ and $\mathbf{V}_\mathbf{F}$ be the eigenvector matrices of \mathbf{E} and \mathbf{F} , respectively. Then one shows that

$$\mathbf{V} = \mathbf{P}_0 \tilde{\mathbf{H}} \mathbf{\Omega} \begin{pmatrix} \mathbf{V}_\mathbf{E} & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_\mathbf{F} \end{pmatrix}$$

is an eigenvector matrix of \mathbf{A} . Note that \mathbf{A} and \mathbf{V} need more storage than $\hat{\mathbf{A}}$, $\mathbf{V}_\mathbf{E}$, and $\mathbf{V}_\mathbf{F}$. The transformations \mathbf{P}_0 , $\tilde{\mathbf{H}}$ and $\mathbf{\Omega}$ follow from Π_0 which can be stored as a vector. Due to the small size the blocks of $\tilde{\mathbf{H}}$ were given explicitly as dense matrices. For larger problems one would prefer the usual sparse form as a rank one update of the identity given in (6).

3 Weighted Equitable Partitions

3.1 Preliminaries

In this section we generalize equitable partitions and accordingly the proposed block triangularization method for square matrices. We introduce the generalized quotient defined for arbitrary partitions of a matrix as a generalization of front and rear quotient. We also introduce the deviation vectors and the deviation matrix and utilize the norm of the latter in order to quantify deviations of a given partition from our generalized notion of equitability. The generalization of the efficient unitary similarity transformation introduced above yields a block triangularization up to an error term due to the deviation from equitability. A further generalization applicable to rectangular matrices preserving only singular values but in general not the spectrum is discussed in the appendix. Note that whenever we invert a matrix explicitly (i.e. not by complex conjugated transposition) this matrix is diagonal. The occasional uses of the pseudo inverse with the property

$$c^\dagger = \begin{cases} 0 & , c = 0 \\ \frac{1}{c} & , \text{else} \end{cases} \quad , \quad c \in \mathbb{C} \quad (13)$$

may be regarded as merely technical.

3.2 Complex Householder Transformations

This subsection aims at the transformation in definition (3) and its properties given in (22). We consider elementary unitary matrices (EUMs) which are rank (at most) one updates of the identity and necessarily (in order to be unitary) [21] of the form

$$\mathbf{U}(\gamma, \mathbf{y}) = \mathbf{I} - \frac{2}{1 + i\gamma} (\mathbf{y}'\mathbf{y})^\dagger \mathbf{y}\mathbf{y}', \quad \mathbf{y} \in \mathbb{C}^n, \gamma \in \mathbb{R}. \quad (14)$$

EUMs are a complex generalization of real Householder matrices [16], [19]. We observe that for $c \in \mathbb{C} \setminus \{0\}$ and \mathbf{P} being a permutation matrix

$$\mathbf{U}(\gamma, c\mathbf{y}) = \mathbf{U}(\gamma, \mathbf{y}) \quad , \quad \mathbf{U}(\gamma, \mathbf{P}\mathbf{y}) = \mathbf{P}\mathbf{U}(\gamma, \mathbf{y})\mathbf{P}'. \quad (15)$$

Let \mathbf{x} and \mathbf{z} be non vanishing complex vectors. We seek an EUM mapping \mathbf{z} into the direction of \mathbf{x} , i.e. a complex vector \mathbf{y} and a real number γ s.t.

$$\mathbf{U}(\gamma, \mathbf{y})\mathbf{z} = \alpha\mathbf{x} \quad \text{with} \quad \alpha \in \mathbb{C} \setminus \{0\}, \quad (16)$$

which implies that $\|\mathbf{x}\|$ and $\|\mathbf{z}\|$ determine α up to a phase factor

$$\sqrt{(\mathbf{U}(\gamma, \mathbf{y})\mathbf{z})'(\mathbf{U}(\gamma, \mathbf{y})\mathbf{z})} = \|\mathbf{z}\| = |\alpha| \|\mathbf{x}\|. \quad (17)$$

Again by (16) \mathbf{y} is a linear combination of \mathbf{x} and \mathbf{z} , namely

$$\mathbf{z} - \alpha \mathbf{x} = \frac{2}{1 + i\gamma} (\mathbf{y}' \mathbf{y})^\dagger (\mathbf{y}' \mathbf{z}) \mathbf{y}. \quad (18)$$

Since according to (15) scaling of \mathbf{y} does not change $\mathbf{U}(\gamma, \mathbf{y})$ we may choose $\mathbf{y} = \mathbf{x} - \frac{1}{\alpha} \mathbf{z}$. We are particular interested in the case $\mathbf{z} = \mathbf{f}_n$. Setting $\alpha = \beta |\alpha|$ and using (18), we reach in the non trivial case, $\mathbf{y} \neq 0$,

$$\gamma(\mathbf{x}, \beta) = (\|\mathbf{x}\| - \operatorname{Re}(\beta x^1))^\dagger \operatorname{Im}(\beta x^1). \quad (19)$$

Thus, the required EUM of \mathbf{f}_n into the direction of \mathbf{x} is determined up to a complex parameter β lying on the unit circle. We introduce

$$\mathbf{H}(\mathbf{x}, \beta) = \mathbf{U}(\gamma(\mathbf{x}, \beta), \mathbf{x} - \|\mathbf{x}\| \bar{\beta} \mathbf{f}_n) \quad \text{with } |\beta| = 1 \quad (20)$$

and give an explicit definition.

Definition 3. Let $\mathbf{x} \in \mathbb{C}^n$ with $\|\mathbf{x}\| > 0$, let $x^1 = \mathbf{f}_n' \mathbf{x}$ denote its first entry and let β be a complex number with $|\beta| = 1$, then

$$\mathbf{H}(\mathbf{x}, \beta) = \begin{cases} \mathbf{I}_n & , \frac{1}{\|\mathbf{x}\|} \mathbf{x} = \bar{\beta} \mathbf{f}_n \\ \mathbf{I}_n + \frac{(\mathbf{x} - \|\mathbf{x}\| \bar{\beta} \mathbf{f}_n)(\mathbf{x} - \|\mathbf{x}\| \bar{\beta} \mathbf{f}_n)'}{\|\mathbf{x}\| \bar{\beta} (x^1 - \|\mathbf{x}\| \beta)} & , \text{ else.} \end{cases}$$

Using $\mathbf{y} = \mathbf{x} - \|\mathbf{x}\| \bar{\beta} \mathbf{f}_n$ we may rewrite

$$\mathbf{H}(\mathbf{x}, \beta) = \mathbf{I}_n + \frac{\beta}{\|\mathbf{x}\|} (\mathbf{y}' \mathbf{f}_n)^\dagger \mathbf{y} \mathbf{y}' = \mathbf{I}_n - (\mathbf{x}' \mathbf{y})^\dagger \mathbf{y} \mathbf{y}'. \quad (21)$$

And we summarize the following properties

$$\mathbf{H}(\mathbf{x}, \beta) \mathbf{f}_n = \frac{\beta}{\|\mathbf{x}\|} \mathbf{x} \quad \text{and} \quad \mathbf{H}(\mathbf{x}, \beta)' \mathbf{x} = \frac{\|\mathbf{x}\|}{\beta} \mathbf{f}_n. \quad (22)$$

Since $\mathbf{H}(\mathbf{x}, \beta)$ is a rank one update of \mathbf{I}_n , it can be stored with $O(n)$ and multiplied with a square matrix of size n in $O(n^2)$. Note that $\mathbf{H}(\mathbf{x}, \beta)$ crucially depends on the ordering of the entries of \mathbf{x} ,

$$\mathbf{H}(\mathbf{P}' \mathbf{x}, \beta) \neq \mathbf{P}' \mathbf{H}(\mathbf{x}, \beta) \mathbf{P} \quad \text{for general } \mathbf{x} \text{ and permutation matrix } \mathbf{P}. \quad (23)$$

Although its norm is determined to be 1, the actual choice of β is arbitrary. We may exploit that freedom in order to enhance the numerical properties of $\mathbf{H}(\beta, \mathbf{x})$. Particular useful is a choice s.t. $\beta x^1 \in \mathbb{R}$, implying $\gamma = 0$ by (19) and leading to a Hermitian matrix. Furthermore, for real \mathbf{x} , $\beta \in \{-1, 1\}$ ensures a real matrix. A practical recommendation might be

$$\textbf{Definition 4. } \beta_0(\mathbf{x}) = \begin{cases} -\frac{\bar{x}^1}{|x^1|} & , x^1 \neq 0 \\ 1 & , x^1 = 0 \end{cases}, \quad \mathbf{x} \in \mathbb{C}^n,$$

which supports numerical stability and coincides with the usual recommendation for the numerical construction of a real Householder matrix. In the previous section we applied β_0 tacitly.

3.3 Weighted Partition, Quotient and Deviation Matrix

Let $\Pi = (c_1, \dots, c_k)$ be a partition of $\{1, \dots, N\}$ into k cells with indicator matrix \mathbf{B} . Let $\mathbf{w} \in \mathbb{C}^N$ and $\mathbf{A} \in \mathbb{C}^{N \times N}$. We introduce the *weighted indicator matrix* $\mathbf{W} = \text{diag}(\mathbf{w}) \mathbf{B}$.

Definition 5. Let $\Pi = (c_1, \dots, c_k)$ be a partition of $\{1, \dots, N\}$. Let $\mathbf{w} \in \mathbb{C}^N$ and let w^v denote its v -th entry.

$$\mathbf{W} = (w_{vi}) \in \mathbb{C}^{N \times k} \text{ with } w_{vi} = \begin{cases} w^v & v \in c_i \\ 0 & , \text{ else} \end{cases}$$

A weighted indicator matrix \mathbf{W} is called *admissible* if $\|\mathbf{w}_i\|$ for all vector blocks \mathbf{w}_i induced by c_i . This implies that $\mathbf{W}'\mathbf{W}$ is invertible and ultimately ensures that the complete spectrum of the quotient, to be defined below, is contained in the spectrum of \mathbf{A} . For the rest of this section we assume admissibility.

We call \mathbf{W} *suitably indexed* if for $u \in c_i, v \in c_j$ it holds that $i < j$ implies $u < v$. In that case the index set is ordered block wise and \mathbf{W} is explicitly block diagonal. In order to simplify the exposition, we may w.l.o.g. assume a suitable indexing.

Definition 6. Let $\mathbf{A} \in \mathbb{C}^{N \times N}$ and let \mathbf{W} be an admissible weighted indicator matrix and let $\alpha \in \mathbb{R}$. The generalized quotient \mathbf{E}^α is given by

$$\mathbf{E}^\alpha = (\mathbf{W}'\mathbf{W})^{-\frac{1-\alpha}{2}} \mathbf{W}'\mathbf{A}\mathbf{W} (\mathbf{W}'\mathbf{W})^{-\frac{1+\alpha}{2}}.$$

We call \mathbf{E}^0 the *Rayleigh quotient*. The matrix entries of \mathbf{E}^α are

$$e_{ij}^\alpha = \left(\frac{1}{\|\mathbf{w}_i\|} \right)^{(1-\alpha)} \mathbf{w}_i' \mathbf{A}_{ij} \mathbf{w}_j \left(\frac{1}{\|\mathbf{w}_j\|} \right)^{(1+\alpha)}. \quad (24)$$

Since $\mathbf{E}^\alpha = (\mathbf{W}\mathbf{W})^{\frac{\alpha}{2}} \mathbf{E}^0 (\mathbf{W}\mathbf{W})^{-\frac{\alpha}{2}}$, all generalized quotients are similar. We distinguish the *front quotient* $\mathbf{E}^- = \mathbf{E}^{-1}$ and the *rear quotient* $\mathbf{E}^+ = \mathbf{E}^1$. The matrix \mathbf{A} is called *front equitable* w.r.t. \mathbf{W} if and only if

$$\mathbf{A}\mathbf{W} = \mathbf{W}\mathbf{E}^- \quad , \text{ i.e. } \quad \forall i, j \in \{1, \dots, k\} \quad \mathbf{A}_{ij} \mathbf{w}_j = e_{ij}^- \mathbf{w}_i \quad (25)$$

and we call \mathbf{A} *rear equitable* w.r.t. \mathbf{W} if and only if

$$\mathbf{W}'\mathbf{A} = \mathbf{E}^+ \mathbf{W}' \quad , \text{ i.e. } \quad \forall i, j \in \{1, \dots, k\} \quad \mathbf{w}_i' \mathbf{A}_{ij} = e_{ij}^+ \mathbf{w}_j'. \quad (26)$$

Definition 7. Maintaining the notation above the front and rear deviation vectors are defined respectively as

$$\mathbf{t}_{ij}^- = \frac{1}{\|\mathbf{w}_j\|} \left(\mathbf{A}_{ij} \mathbf{w}_j - e_{ij}^- \mathbf{w}_i \right) \quad \text{and} \quad \mathbf{t}_{ij}^+ = \frac{1}{\|\mathbf{w}_i\|} \left(\mathbf{w}_i' \mathbf{A}_{ij} - e_{ij}^+ \mathbf{w}_j' \right)',$$

and the front and rear deviation matrices are

$$\mathbf{T}^\pm = \begin{pmatrix} \mathbf{t}_{11}^\pm & \cdots & \mathbf{t}_{1k}^\pm \\ \vdots & \ddots & \vdots \\ \mathbf{t}_{k1}^\pm & \cdots & \mathbf{t}_{kk}^\pm \end{pmatrix} \in \mathbb{C}^{N \times k}, \text{ i.e. } \quad \begin{aligned} \mathbf{T}^- &= (\mathbf{A}\mathbf{W} - \mathbf{W}\mathbf{E}^-)(\mathbf{W}'\mathbf{W})^{-\frac{1}{2}} \\ \mathbf{T}^+ &= (\mathbf{A}'\mathbf{W} - \mathbf{W}\mathbf{E}^{+'})(\mathbf{W}'\mathbf{W})^{-\frac{1}{2}}. \end{aligned}$$

The entries of \mathbf{E}^\pm and the deviation vectors have an intuitive interpretation in the framework of ordinary equitability arising for $\mathbf{w}_i = \mathbf{j}_{n_i}$. Then e_{ij}^- and $\|\mathbf{t}_{ij}^-\|$ (e_{ij}^+ and $\|\mathbf{t}_{ij}^+\|$) are the mean and the standard deviation of the row (column) sums of \mathbf{A}_{ij} .

Scaling the vector blocks \mathbf{w}_i by $\mu_i \in \mathbb{C} \setminus \{0\}$ changes the entries of the generalized quotient to $\mu_i^\alpha e_{ij}^\alpha \mu_j^{-\alpha}$ although such a transformation sustains equitability (if present). Note that e_{ij}^0 and $\|\mathbf{t}_{ij}^\pm\|$, and therefore the singular values of \mathbf{T}^\pm , are independent of such a scaling. By definition, \mathbf{T}^\pm is an all zero matrix if and only if its respective equitability holds. At the end of this section, we will consider suitable norms of \mathbf{T}^\pm as measures for deviation from equitability.

3.4 (Approximate) Block Triangularization

Let \mathbf{W} be an admissible weighted indicator matrix of a partition $\Pi = (c_1, \dots, c_k)$ with weight vector $\mathbf{w} \in \mathbb{C}^N$ and indicator matrix \mathbf{B} . Let \mathbf{w}_i be induced by c_i . Replacing \mathbf{w}_i by \mathbf{f}_{n_i} for all i yields the new vector \mathbf{f} . Let $\mathbf{N} = (\mathbf{W}'\mathbf{W})^{\frac{1}{2}}$ and let $\mathbf{V} = \text{diag}(\beta_1, \dots, \beta_k)$ be a unitary diagonal matrix of size k . We introduce

$$\mathbf{Y}(\mathbf{W}, \mathbf{V}) = \mathbf{Y}(\mathbf{w}, \Pi, \mathbf{V}) = \text{diag}(\mathbf{w})\mathbf{B} - \text{diag}(\mathbf{f})\mathbf{B}\mathbf{N}\mathbf{V}', \quad (27)$$

which has the form of a weighted indicator matrix. The actual choice of \mathbf{V} is a priori arbitrary. This freedom may be exploited in order to enhance the numerical properties of the transformation matrix given in the next definition.

Definition 8. Let \mathbf{Y} be derived from an admissible weighted indicator matrix \mathbf{W} and a unitary diagonal matrix \mathbf{V} as above, then

$$\mathbf{H}(\mathbf{W}, \mathbf{V}) = \mathbf{I}_N - \mathbf{Y}(\mathbf{W}'\mathbf{Y})^\dagger \mathbf{Y}'.$$

Since \mathbf{Y} and \mathbf{W} have the same block diagonal form, $\mathbf{Y}'\mathbf{W}$ is a diagonal matrix and $\mathbf{H}(\mathbf{V}, \mathbf{W})$ is block diagonal, hence its numerical properties are comparable to those of a single Householder matrix. In particular, the costs for computing and storing are of order $O(N)$, and it can be applied to a square matrix in $O(N^2)$. For suitably indexed \mathbf{W} the block diagonal form of $\mathbf{H}(\mathbf{W}, \mathbf{V})$ is explicit,

$$\mathbf{H}(\mathbf{W}, \mathbf{V}) = \text{diag}(\mathbf{H}(\mathbf{w}_1, \beta_1), \dots, \mathbf{H}(\mathbf{w}_k, \beta_k)). \quad (28)$$

The diagonal blocks are given in definition (3). For $\mathbf{A} \in \mathbb{C}^{N \times N}$ we consider

$$\tilde{\mathbf{A}} = \mathbf{H}(\mathbf{W}, \mathbf{V})' \mathbf{A} \mathbf{H}(\mathbf{W}, \mathbf{V}) \quad \text{with} \quad \tilde{\mathbf{A}}_{ij} = \mathbf{H}(\mathbf{w}_i, \beta_i)' \mathbf{A}_{ij} \mathbf{H}(\mathbf{w}_j, \beta_j). \quad (29)$$

By the properties (22) of the $\mathbf{H}(\mathbf{w}_i, \beta_i)$ it follows that

$$\tilde{\mathbf{A}}_{ij} \mathbf{f}_{n_j} \sim \mathbf{f}_{n_i} \quad \forall i, j \text{ if and only if } \mathbf{A} \text{ is front equitable w.r.t. } \mathbf{W}, \quad (30)$$

$$\mathbf{f}'_{n_i} \tilde{\mathbf{A}}_{ij} \sim \mathbf{f}'_{n_j} \quad \forall i, j \text{ if and only if } \mathbf{A} \text{ is rear equitable w.r.t. } \mathbf{W}. \quad (31)$$

If we consider for a moment front (row) equitability, the first column (row) of each block $\tilde{\mathbf{A}}_{ij}$ would be all zero from its second to last entry. This implies an implicit block triangular form of $\tilde{\mathbf{A}}$, which can be made explicit by the following permutation mapping the first index of each cell accordingly into $\{1, \dots, k\}$.

Definition 9. Let $\mathbf{n} = (n_1, \dots, n_k)$ be a sequence of k positive integers with $\sum_{i=1}^k n_i = N$. The permutation $\Omega_{\mathbf{n}} : \{1, \dots, N\} \rightarrow \{1, \dots, N\}$ is defined by

$$\Omega_{\mathbf{n}} \left(m_i + \sum_{j=i}^{i-1} n_j \right) = \begin{cases} i & , m_i = 1 \\ k - i + m_i + \sum_{j=i}^{i-1} n_j & , m_i \in \{2, \dots, n_i\} \end{cases}$$

with $i \in \{1, \dots, k\}$.

We proceed with the general case and give the following theorem, which may be seen as a corollary of theorem (2). In order to keep this section self contained, it is proven independently.

Theorem 1. Let $\Pi = (c_1, \dots, c_k)$ be an admissible partition for $\mathbf{A} \in \mathbb{C}^{N \times N}$ and $\mathbf{w} \in \mathbb{C}^N$ with weighted indicator matrix $\mathbf{W} \in \mathbb{C}^{N \times k}$, generalized quotient \mathbf{E}^α and deviation matrices \mathbf{T}^\pm . Let $\mathbf{V} = \text{diag}(\beta_1, \dots, \beta_k)$ be a unitary diagonal matrix and let $\tilde{\mathbf{H}} = \mathbf{H}(\mathbf{W}, \mathbf{V})$ as in definition (8) and let Ω be the permutation matrix corresponding to $\Omega_{(|c_1|, \dots, |c_k|)}$. Let

$$\tilde{\mathbf{A}} = \tilde{\mathbf{H}}' \mathbf{A} \tilde{\mathbf{H}} = \begin{pmatrix} \tilde{\mathbf{A}}_{11} & \cdots & \tilde{\mathbf{A}}_{1k} \\ \vdots & \ddots & \vdots \\ \tilde{\mathbf{A}}_{k1} & \cdots & \tilde{\mathbf{A}}_{kk} \end{pmatrix}, \quad \tilde{\mathbf{A}}_{ij} = \mathbf{H}(\mathbf{w}_i, \beta_i)' \mathbf{A}_{ij} \mathbf{H}(\mathbf{w}_j, \beta_j)$$

and

$$\hat{\mathbf{A}} = \Omega' \tilde{\mathbf{A}} \Omega = \begin{pmatrix} \mathbf{E} & \mathbf{D}^{+'} \\ \mathbf{D}^- & \mathbf{F} \end{pmatrix} \quad \text{with } \mathbf{E} \in \mathbb{C}^{k \times k}.$$

Then $\hat{\mathbf{A}}$ is unitarily similar to \mathbf{A} , the upper left block \mathbf{E} is unitarily similar to the Rayleigh quotient \mathbf{E}^0 and the off-diagonal blocks \mathbf{D}^\pm have the same singular values as \mathbf{T}^\pm , respectively. Additionally, any eigenvector $\hat{\mathbf{z}}$ of $\hat{\mathbf{A}}$ yields an eigenvector $\mathbf{z} = \tilde{\mathbf{H}} \Omega \hat{\mathbf{z}}$ of \mathbf{A} to the same eigenvalue.

Proof. Unitary similarity to \mathbf{A} follows from the unitarity of $\tilde{\mathbf{H}}$ and Ω .

Considering the matrix blocks $\tilde{\mathbf{A}}_{ij}$ of $\tilde{\mathbf{A}}$ induced by cells c_i and c_j we have

$$e_{ij} = \mathbf{f}'_{n_i} \tilde{\mathbf{A}}_{ij} \mathbf{f}_{n_j} = \frac{\overline{\beta_i}}{\|\mathbf{w}_i\|} \frac{\beta_j}{\|\mathbf{w}_j\|} \mathbf{w}_i' \mathbf{A}_{ij} \mathbf{w}_j = \frac{\beta_j}{\beta_i} e_{ij}^0. \quad (32)$$

By $\mathbf{\Omega}$ those e_{ij} are mapped accordingly into the upper left block \mathbf{E} . Therefore, we may rewrite $\mathbf{E} = \mathbf{V}'\mathbf{E}^0\mathbf{V}$, which proofs unitary similarity of \mathbf{E} and \mathbf{E}^0 . In order to show that \mathbf{D}^\pm is unitarily equivalent to \mathbf{T}^\pm , we observe that by the properties of $\mathbf{\Omega}$ we can write \mathbf{D}^\pm as

$$\mathbf{D}^\pm = \begin{pmatrix} \mathbf{d}_{11}^\pm & \cdots & \mathbf{d}_{1k}^\pm \\ \vdots & \ddots & \vdots \\ \mathbf{d}_{k1}^\pm & \cdots & \mathbf{d}_{kk}^\pm \end{pmatrix} \in \mathbb{C}^{(N-k) \times k}, \quad (33)$$

wherein \mathbf{d}_{ij}^- is the first column and $\mathbf{d}_{ij}^{+'}$ is the first row of the matrix block $\tilde{\mathbf{A}}_{ij}$ starting from the second entry. We have

$$\begin{pmatrix} 0 \\ \mathbf{d}_{ij}^- \end{pmatrix} = \tilde{\mathbf{A}}_{ij}\mathbf{f}_{n_j} - e_{ij}^-\mathbf{f}_{n_i} = \mathbf{H}(\beta_i, \mathbf{w}_i)' \mathbf{t}_{ij}^-, \quad (34)$$

$$(0, \mathbf{d}_{ij}^{+'}) = \mathbf{f}_{n_i}' \tilde{\mathbf{A}}_{ij} - e_{ij}^+\mathbf{f}_{n_j}' = \mathbf{t}_{ij}^{+'} \mathbf{H}(\beta_j, \mathbf{w}_j), \quad (35)$$

which shows that

$$\mathbf{\Omega}'\tilde{\mathbf{H}}'\mathbf{T}^\pm = \begin{pmatrix} \mathbf{0} \\ \mathbf{D}^\pm \end{pmatrix}. \quad (36)$$

The eigenvector relation can be shown by applying $\tilde{\mathbf{H}}\mathbf{\Omega}$ from the left to

$$\lambda\hat{\mathbf{z}} = \hat{\mathbf{A}}\hat{\mathbf{z}} = \mathbf{\Omega}'\tilde{\mathbf{H}}'\mathbf{A}\mathbf{z}. \quad (37)$$

□

3.5 Deviation from Equitability

Let $\|\cdot\|_U$ denote a unitarily invariant norm.

Corollary 1.

$$\|\mathbf{D}^\pm\|_U = \|\mathbf{T}^\pm\|_U.$$

Corollary 2. Let $\mathbf{T}_\Theta^- = (\mathbf{A}\mathbf{W} - \mathbf{W}\Theta)\mathbf{N}^{-1}$ and $\mathbf{T}_\Theta^+ = \mathbf{N}^{-1}(\mathbf{W}'\mathbf{A} - \Theta\mathbf{W}')$ with $\mathbf{N} = (\mathbf{W}'\mathbf{W})^{\frac{1}{2}}$. Then

$$\|\mathbf{T}^\pm\|_U = \min_{\Theta} \|\mathbf{T}_\Theta^\pm\|_U.$$

The minimum is unique if $\|\cdot\|_U$ is a Schatten norm.

Proof. Applying $\mathbf{\Omega}'\tilde{\mathbf{H}}'$ from the left and \mathbf{V}' from the right to \mathbf{T}_Θ^- yields

$$\begin{aligned} \|\mathbf{\Omega}'\tilde{\mathbf{H}}'\mathbf{T}_\Theta^-\mathbf{V}'\|_U &= \|\bar{\mathbf{A}}\mathbf{\Omega}'\tilde{\mathbf{H}}'\mathbf{W}\mathbf{N}^{-1}\mathbf{V} - \mathbf{\Omega}'\tilde{\mathbf{H}}'\mathbf{W}\Theta\mathbf{N}^{-1}\mathbf{V}\|_U \\ &= \left\| \begin{pmatrix} \mathbf{E} \\ \mathbf{D}^- \end{pmatrix} - \begin{pmatrix} \mathbf{V}'\mathbf{N}\Theta\mathbf{N}^{-1}\mathbf{V} \\ \mathbf{0} \end{pmatrix} \right\|_U \end{aligned} \quad (38)$$

using $\mathbf{\Omega}'\tilde{\mathbf{H}}'\mathbf{W} = \begin{pmatrix} \mathbf{V}'\mathbf{N} \\ \mathbf{0} \end{pmatrix}$. The last term is readily minimized for

$$\Theta = \mathbf{N}^{-1}\mathbf{V}\mathbf{E}\mathbf{V}'\mathbf{N} = (\mathbf{W}'\mathbf{W})^{-\frac{1}{2}}\mathbf{E}^0(\mathbf{W}'\mathbf{W})^{\frac{1}{2}} = \mathbf{E}^-. \quad (39)$$

Obviously, the minimization is unique for several $\|\cdot\|_U$ including the Schatten norms. A similar proof applies for \mathbf{T}^+ . \square

The idea underlying the last proof is essentially the same as in [8, proof of theorem 11]. A particular useful choice for $\|\cdot\|_U$ might be the Frobenius norm, which upper bounds the spectral norm. Its square is simply the sum of the squared norms of the deviation vectors. One may also think of other characterizations for approximate equitable partitions which have moderate computational costs, for instance the number of nonzero columns of \mathbf{T}^\pm , which upper bounds the rank.

4 Discussion and Remarks

4.1 Relating Equitability Deviation and Spectral Deviation

Since $\hat{\mathbf{A}}$ and \mathbf{A} are unitarily similar and by corollaries (1) and (2) of theorem (1), we may in a sense 'measure' the deviation of a partition from being equitable by using a suitable unitarily invariant norm of \mathbf{T}^\pm , yielding a norm of \mathbf{D}^\pm , which in turn may serve as a measure for the deviation of the joint eigenvalue sets or the joint singular value sets of \mathbf{E} and \mathbf{F} from the respective values of \mathbf{A} .

As an example we consider the spectral norm and the eigenvalue bound of Weyl for Hermitian matrices. Assuming Hermiticity we may set $\mathbf{D}^\pm = \mathbf{D}$ and

$$\hat{\mathbf{A}} = \begin{pmatrix} \mathbf{E} & \mathbf{0} \\ \mathbf{0} & \mathbf{F} \end{pmatrix} + \begin{pmatrix} \mathbf{0} & \mathbf{D}' \\ \mathbf{D} & \mathbf{0} \end{pmatrix}. \quad (40)$$

Let $\mu_1 \leq \dots \leq \mu_N$ be the joint spectrum of Hermitian \mathbf{E} and \mathbf{F} , $\lambda_1 \leq \dots \leq \lambda_N$ the eigenvalues of \mathbf{A} and let τ_{spec} be the largest singular value of \mathbf{D} . We have

$$|\mu_i - \lambda_i| \leq \tau_{\text{spec}}, \quad 1 \leq i \leq N \quad (41)$$

by the Weyl inequalities. Many more perturbation bounds on eigenvalues and singular values and their corresponding vectors are feasible, e.g. [10],[9],[11].

4.2 Cognate Concepts

The notion of quasi-block-stochastic matrices of Kuich [18] as a generalization of quasi-stochastic matrices [14] bears a close resemblance to our notion of equitability. A minor difference is that for quasi-block-stochastic matrices it is required that the first entry of each \mathbf{v}_i has to be 1. Kuich also describes how

to exploit this structure to triangularize a (real) matrix by a (real, in general not unitary) similarity transformation using a theorem of Haynsworth [15]. Another similar but less general concept is used by Fiol and Carriga and is called *pseudo-regular* partitions. It considers a positive eigenvector \mathbf{v} of binary matrices [12, pp. 278/9]. The partition Π of the matrix is pseudo-regular if \mathbf{v} and Π induce a (weighted) equitable partition. Since \mathbf{v} is fixed up to a positive scale factor, the pseudo-quotient (i.e. front divisor) is unique.

There are some more techniques in network analysis which can be described as variations of (1) and which are used to partition the node set of a graph (=assigning roles) according to structural properties and to derive a smaller graph (the quotient or image graph) which gives a condensed representation of essential relations between the cells (=roles) of that partition. Some of those are without apparent regard to the spectrum. For instance, Kate and Ravindran introduced *epsilon equitable* partitions for (an adjacency matrix \mathbf{A} of) a simple graph [17]. Let $\Pi = (c_1, \dots, c_k)$ be a partition of the node set of \mathbf{A} . Let \mathbf{A}_{ij} be induced by the i -th row cell and the j -th column cell. Let $\mathbf{r}_{ij} = \mathbf{A}_{ij}\mathbf{j}_{n_j}$ be a column vector of length $n_i = |c_i|$. If

$$\forall i, j \in \{1, \dots, k\} \quad \max_{1 \leq v, w \leq n_i} |\mathbf{r}_{ij,v} - \mathbf{r}_{ij,w}| \leq \epsilon \quad (42)$$

then Π is called ϵ -equitable. The ordinary equitable partition arises for $\epsilon = 0$. Another variation of (1) can be employed to describe the concept of *regular equivalence* [3], which is defined by the restriction that for a partition Π any vector $\mathbf{r}_{ij} = \mathbf{A}_{ij}\mathbf{j}_{n_j}$ must have either no zero entry or all entries zero i.e.

$$\forall i, j \in \{1, \dots, k\} \quad \prod_v \mathbf{r}_{ij,v} = 0 \Rightarrow \sum_v |\mathbf{r}_{ij,v}| = 0. \quad (43)$$

4.3 Finding Equitable Partitions

There are several algorithms for finding ordinary equitable partitions of graphs and matrices, for instance [1], [13].

We sketch the most often employed top-down approach made suitable to the case of finding an ordinary front equitable partition of a complex matrix \mathbf{A} . At each step one considers a temporary partition (initially often the single cell partition) and (sequentially) subdivides any cell c_i for any j according to the entries of $\mathbf{A}_{ij}\mathbf{j}_{n_j}$, called colors, s.t. each subcell is induced by a unique color, until this subdivision is non trivial, resulting in a refined partition. One iterates until any feasible subdivision is trivial, i.e. the final partition is the unique coarsest front equitable refinement (w.r.t. to the initial partition). Of course, this can be adapted for the weighted case. However, if the weight vector \mathbf{w} has no zero entries one may employ the sketched procedure for the unweighted case readily by considering the matrix $\text{diag}(\mathbf{w})^{-1} \mathbf{A} \text{diag}(\mathbf{w})$. This follows by left multiplication of $\text{diag}(\mathbf{w})^{-1}$ to the equitability condition

$$\mathbf{A} \text{diag}(\mathbf{w}) \mathbf{B} = \text{diag}(\mathbf{w}) \mathbf{B} \mathbf{E}^-. \quad (44)$$

In general, the choice of a weight vector \mathbf{w} may be guided by insights into the problem underlying the considered matrix \mathbf{A} . In search for \mathbf{w} , one may also exploit that the columns of the weighted indicator matrix \mathbf{W} are a basis for the linear span of all eigenvectors of \mathbf{A} corresponding to eigensolutions of \mathbf{E}^- . As an example, let \mathbf{x} and \mathbf{y} be two such eigenvectors for different eigenvalues. Considering them separately using the top down approach above one finds the single cell partition since \mathbf{x} and \mathbf{y} are eigenvectors. This may be avoided by using a non trivial linear combination, which lies in the linear span of the columns of \mathbf{W} but is not an eigenvector.

How to find partitions with suitably small but non zero deviation from equitability is out of the scope of this article.

A Generalization as a Singular Value Decomposition

Our proposed method for block triangularization can be described as an employment of a one-step singular value decomposition (SVD) of the weighted indicator matrix \mathbf{W} as

$$\mathbf{W} = \begin{bmatrix} \tilde{\mathbf{H}}\mathbf{\Omega} \\ \mathbf{0} \end{bmatrix} \begin{pmatrix} \mathbf{N} \\ \mathbf{0} \end{pmatrix} \mathbf{V}' \quad (45)$$

wherein the square diagonal matrix \mathbf{N} contains the singular values of \mathbf{W} and $\mathbf{V} = \text{diag}(\beta_1, \dots, \beta_k)$ is unitary diagonal. In deed, if we interpret \mathbf{f}_{n_i} and the vector blocks \mathbf{w}_i as matrices in $\mathbb{C}^{n_i \times 1}$, then \mathbf{w}_i has the SVD

$$\mathbf{w}_i = \mathbf{H}(\mathbf{w}_i, \beta_i) \frac{\|\mathbf{w}_i\|}{\beta_i} \mathbf{f}_{n_i} = \mathbf{H}(\mathbf{w}_i, \beta_i) \begin{pmatrix} \|\mathbf{w}_i\| \\ \mathbf{0} \end{pmatrix} \overline{\beta_i}. \quad (46)$$

In that view, one may obtain a generalization by replacing the non vanishing vector blocks \mathbf{w}_i by rectangular matrix blocks \mathbf{W}_i with maximal column rank. In the remainder of this section we build on this idea and derive an approximate block triangularization of a rectangular matrix $\mathbf{A} \in \mathbb{C}^{m \times n}$, using given SVDs of a pair of block diagonal matrices with maximal column rank, acting on the rows and columns of \mathbf{A} respectively and separately.

For notational convenience we define a 2×1 block matrix with empty lower block and the identity matrix in the square upper block.

Definition 10. Let r and n be positive integers with $r \leq n$.

$$\mathbf{I}_n^r = \begin{pmatrix} \mathbf{I}_r \\ \mathbf{0} \end{pmatrix} \in \{0, 1\}^{n \times r}.$$

We may identify $\mathbf{I}_n^1 = \mathbf{f}_n$. As a block diagonal generalization we define

Definition 11. Let $\mathbf{r} = (r_1, \dots, r_k)$ and $\mathbf{n} = (n_1, \dots, n_k)$ be ordered sequences of positive integers, s.t. $r_i \leq n_i$.

$$\mathbf{I}_{\mathbf{n}}^{\mathbf{r}} = \text{diag}(\mathbf{I}_{n_1}^{r_1}, \dots, \mathbf{I}_{n_k}^{r_k}).$$

We will also utilize the following permutation.

Definition 12. Let $\mathbf{r} = (r_1, \dots, r_k)$ and $\mathbf{n} = (n_1, \dots, n_k)$ be ordered sequences of positive integers s.t. $\forall i \in \{1, \dots, k\} r_i \leq n_i$. Let $r = \sum_i r_i$ and $n = \sum_i n_i$. Then $\Omega_{\mathbf{n}}^{\mathbf{r}} : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ is defined by

$$\Omega_{\mathbf{n}}^{\mathbf{r}} \left(s_i + \sum_{j=1}^{i-1} n_j \right) = \begin{cases} s_i + \sum_{j=1}^{i-1} r_j & , 0 < s_i \leq r_i \\ s_i + r + \sum_{j=1}^{i-1} (n_j - r_j) & , r_i < s_i \leq n_i \end{cases}$$

with $i \in \{1, \dots, k\}$.

$\Omega_{\mathbf{n}}^{\mathbf{r}}$ maps the first r_i elements of cell i into the first r elements.

Proposition 1. In the notation of definitions (11) and (12) above, let Ω be the permutation matrix corresponding to $\Omega_{\mathbf{n}}^{\mathbf{r}}$. Then $\Omega' \mathbf{I}_{\mathbf{n}}^{\mathbf{r}} = \mathbf{I}_n^{\mathbf{r}}$.

Definition 13. Let $\mathbf{W}^- \in \mathbb{C}^{m \times q}$ be a block diagonal matrix with l diagonal blocks $\mathbf{W}_i^- \in \mathbb{C}^{m_i \times q_i}$ of rank q_i and let $\mathbf{W}^+ \in \mathbb{C}^{n \times r}$ be a block diagonal matrix with k diagonal blocks $\mathbf{W}_i^+ \in \mathbb{C}^{n_i \times r_i}$ each of rank r_i . Let singular value decompositions for the \mathbf{W}_i^- be given by

$$\mathbf{W}_i^- = \mathbf{U}_i^- \mathbf{S}_i^- \mathbf{V}_i^{-'} = \mathbf{U}_i^- (\mathbf{I}_{m_i}^{q_i} \mathbf{N}_i^-) \mathbf{V}_i^{-'}$$

with square unitary $\mathbf{U}_i^- \in \mathbb{C}^{m_i \times m_i}$ and $\mathbf{V}_i^- \in \mathbb{C}^{q_i \times q_i}$, and with $\mathbf{S}_i^- = \mathbf{I}_{m_i}^{q_i} \mathbf{N}_i^-$ wherein $\mathbf{N}_i^- \in \mathbb{R}^{q_i \times q_i}$ is a square diagonal matrix with positive diagonal elements. Let $\mathbf{U}^- \in \mathbb{C}^{m \times m}$, $\mathbf{S}^- \in \mathbb{C}^{m \times q}$, $\mathbf{N}^- \in \mathbb{R}^{q \times q}$, and $\mathbf{V}^- \in \mathbb{C}^{q \times q}$ be block diagonal with l diagonal blocks given by \mathbf{U}_i^- , \mathbf{S}_i^- , \mathbf{N}_i^- , and \mathbf{V}_i^- , respectively. Let Ω^- be the permutation matrix corresponding to the permutation $\Omega_{(m_1, \dots, m_l)}^{(q_1, \dots, q_l)}$ s.t. $\Omega^{-'} \mathbf{S}^- = \mathbf{I}_m^q \mathbf{N}^-$. This induces a singular value decomposition of \mathbf{W}^- as

$$\mathbf{W}^- = \mathbf{U}^- (\Omega^{-'} \mathbf{S}^-) \mathbf{V}^{-'} = \mathbf{U}^- (\mathbf{I}_m^q \mathbf{N}^-) \mathbf{V}^{-'}.$$

Let the corresponding relations hold for \mathbf{W}^+ and let $\mathbf{A} \in \mathbb{C}^{m \times n}$. Define the Rayleigh quotient $\mathbf{E}^0 \in \mathbb{C}^{q \times r}$ as a block matrix with

$$\mathbf{E}^0 = \begin{pmatrix} \mathbf{E}_{11}^0 & \cdots & \mathbf{E}_{1n}^0 \\ \vdots & \ddots & \vdots \\ \mathbf{E}_{1m}^0 & \cdots & \mathbf{E}_{mn}^0 \end{pmatrix}, \quad \mathbf{E}_{ij}^0 = \mathbf{V}_i^- \mathbf{I}_{m_i}^{q_i} {}' \mathbf{U}_i^{-'} \mathbf{A}_{ij} \mathbf{U}_j^+ \mathbf{I}_{n_j}^{r_j} \mathbf{V}_j^{+'} \in \mathbb{C}^{q_i \times r_j}.$$

The front and rear deviation matrices, $\mathbf{T}^- \in \mathbb{C}^{m \times r}$ and $\mathbf{T}^+ \in \mathbb{C}^{n \times q}$ respectively, are block matrices with

$$\begin{aligned} \mathbf{T}_{ij}^- &= \mathbf{A}_{ij} \mathbf{U}_j^+ \mathbf{I}_{n_j}^{r_j} \mathbf{V}_j^{+'} - \mathbf{U}_i^- \mathbf{I}_{m_i}^{q_i} \mathbf{V}_i^{-'} \mathbf{E}_{ij}^0 \in \mathbb{C}^{m_i \times r_i}, \\ \mathbf{T}_{ij}^+ &= (\mathbf{A}_{ij})' \mathbf{U}_i^- \mathbf{I}_{m_i}^{q_i} \mathbf{V}_i^{-'} - \mathbf{U}_j^+ \mathbf{I}_{n_j}^{r_j} \mathbf{V}_j^{+'} (\mathbf{E}_{ij}^0)' \in \mathbb{C}^{n_i \times q_i}. \end{aligned}$$

Proposition 2. *In the notation of definition (13), \mathbf{E}^0 , \mathbf{T}^- , and \mathbf{T}^+ are identical for all singular value decompositions of \mathbf{W}^- and \mathbf{W}^+ which obey the block diagonal form.*

Proof. \mathbf{E}^0 and \mathbf{T}^\pm can be entirely expressed in terms of \mathbf{W}^\pm since

$$\mathbf{U}^- \mathbf{I}_m^q \mathbf{V}^{-'} = \mathbf{W}^- \left(\mathbf{W}^{-'} \mathbf{W}^- \right)^{-\frac{1}{2}} \text{ and } \mathbf{U}^+ \mathbf{I}_n^r \mathbf{V}^{+'} = \mathbf{W}^+ \left(\mathbf{W}^{+'} \mathbf{W}^+ \right)^{-\frac{1}{2}}. \quad (47)$$

$$\mathbf{E}^0 = \left(\mathbf{W}^{-'} \mathbf{W}^- \right)^{-\frac{1}{2}} \mathbf{W}^{-'} \mathbf{A} \mathbf{W}^+ \left(\mathbf{W}^{+'} \mathbf{W}^+ \right)^{-\frac{1}{2}}, \quad (48)$$

$$\mathbf{T}^- = \mathbf{A} \mathbf{W}^+ \left(\mathbf{W}^{+'} \mathbf{W}^+ \right)^{-\frac{1}{2}} - \mathbf{W}^{-'} \left(\mathbf{W}^{-'} \mathbf{W}^- \right)^{-\frac{1}{2}} \mathbf{E}^0, \quad (49)$$

$$\mathbf{T}^+ = \mathbf{A}' \mathbf{W}^- \left(\mathbf{W}^{-'} \mathbf{W}^- \right)^{-\frac{1}{2}} - \mathbf{W}^{+'} \left(\mathbf{W}^{+'} \mathbf{W}^+ \right)^{-\frac{1}{2}} (\mathbf{E}^0)'. \quad (50)$$

□

Theorem 2. *In the notation of definition (13) above, let*

$$\hat{\mathbf{A}} = \mathbf{\Omega}^{-'} \mathbf{U}^{-'} \mathbf{A} \mathbf{U}^+ \mathbf{\Omega}^+ = \begin{pmatrix} \mathbf{E} & \mathbf{D}^{+'} \\ \mathbf{D}^- & \mathbf{F} \end{pmatrix} \quad \text{with } \mathbf{E} \in \mathbb{C}^{q \times r}.$$

and let $\|\cdot\|_U$ be a unitarily invariant matrix norm. Then \mathbf{E} and \mathbf{E}^0 are unitarily equivalent, \mathbf{D}^\pm and \mathbf{T}^\pm have the same singular values, respectively, and

$$\begin{aligned} \|\mathbf{T}^-\|_U &= \min_{\mathbf{\Theta}} \|\mathbf{A} \mathbf{U}^+ \mathbf{I}_n^r \mathbf{V}^{+'} - \mathbf{U}^- \mathbf{I}_m^q \mathbf{V}^{-'} \mathbf{\Theta}\|_U \\ \|\mathbf{T}^+\|_U &= \min_{\mathbf{\Theta}} \|\mathbf{A}' \mathbf{U}^- \mathbf{I}_m^q \mathbf{V}^{-'} - \mathbf{U}^+ \mathbf{I}_n^r \mathbf{V}^{+'} \mathbf{\Theta}'\|_U. \end{aligned}$$

Proof. Unitary equivalence of \mathbf{E} and \mathbf{E}^0 follows from

$$\mathbf{E} = \mathbf{I}_m^q \mathbf{\hat{A}} \mathbf{I}_n^r = \text{diag}(\mathbf{U}_1^- \mathbf{I}_{m_1}^{q_1}, \dots, \mathbf{U}_l^- \mathbf{I}_{m_l}^{q_l})' \mathbf{A} \text{diag}(\mathbf{U}_1^+ \mathbf{I}_{n_1}^{r_1}, \dots, \mathbf{U}_k^+ \mathbf{I}_{n_k}^{r_k}), \quad (51)$$

which uses proposition (1), yielding $\mathbf{E}^0 = \mathbf{V}^- \mathbf{E} \mathbf{V}^{+'}$.

That \mathbf{D}^- and \mathbf{T}^- share the same multiset of singular values follows from

$$\begin{aligned} \begin{pmatrix} \mathbf{0} \\ \mathbf{D}^- \end{pmatrix} &= \hat{\mathbf{A}} \mathbf{I}_n^r - \mathbf{I}_m^q \mathbf{E} = \mathbf{\Omega}^{-'} \mathbf{U}^{-'} \mathbf{A} \mathbf{U}^+ \mathbf{\Omega}^+ \mathbf{I}_n^r - \mathbf{I}_m^q \mathbf{V}^{-'} \mathbf{E}^0 \mathbf{V}^+ \\ &= \mathbf{\Omega}^{-'} \mathbf{U}^{-'} \mathbf{T}^- \mathbf{V}^+. \end{aligned} \quad (52)$$

The proof for \mathbf{D}^+ and \mathbf{T}^+ is analogous.

Applying the unitary matrices $\mathbf{\Omega}^{-'} \mathbf{U}^{-'}$ from the left and \mathbf{V}^+ from the right to the second term in the penultimate equation of theorem (2) yields

$$\min_{\mathbf{\Theta}} \|\tilde{\mathbf{A}} \mathbf{\Omega}^{+'} \mathbf{I}_n^r - \mathbf{\Omega}^{-'} \mathbf{I}_m^q \mathbf{V}^{-'} \mathbf{\Theta} \mathbf{V}^+\|_U = \left\| \begin{pmatrix} \mathbf{E} \\ \mathbf{D}^- \end{pmatrix} - \begin{pmatrix} \mathbf{V}^{-'} \mathbf{\Theta} \mathbf{V}^+ \\ \mathbf{0} \end{pmatrix} \right\|_U. \quad (53)$$

The last term is readily minimized for $\mathbf{\Theta} = \mathbf{V}^- \mathbf{E} \mathbf{V}^{+'} = \mathbf{E}^0$. The minimum is obviously unique if $\|\cdot\|_U$ is a Schatten norm. A similar proof applies for the minimum property of \mathbf{T}^+ . □

References

- [1] O. Bastert. Computing equitable partitions of graphs. *MATCH Communications in Mathematical and in Computer Chemistry*, 40:265–272, 1999.
- [2] P. Boldi and S. Vigna. Fibrations of graphs. *Discrete Math*, 243, 2002.
- [3] S. P. Borgatti and M. G. Everett. Regular equivalence: general theory. *Journal of Mathematical Sociology*, 19:29–52, 1994.
- [4] P. Buchholz. Exact and ordinary lumpability in finite markov chains. *Journal of Applied Probability*, 31(1):59–75, 1994.
- [5] J. Cai, M. Führer, and N. Immerman. An optimal lower bound on the number of variables for graphs. *Combinatorica*, 12(4):389–410, 1992.
- [6] T.-J. Chang, B.-S. Tam, and S.-H. Wu. Theorems on partitioned matrices revisited and their applications to graph spectra. *Linear Algebra and its Applications*, 434(2):559 – 581, 2011.
- [7] D. M. Cvetković, M. Doob, and H. Sachs. *Spectra of graphs: theory and application*. Pure and applied mathematics. Academic Press, 1980.
- [8] A. Dax. On extremum properties of orthogonal quotients matrices. *Linear Algebra and its Applications*, 432(5):1234 – 1257, 2010.
- [9] A.S. Deif. Rigorous perturbation bounds for eigenvalues and eigenvectors of a matrix. *Journal of Computational and Applied Mathematics*, 57(3):403 – 412, 1995.
- [10] S. C. Eisenstat and I. C. F. Ipsen. Three absolute perturbation bounds for matrix eigenvalues imply relative bounds. *SIAM Journal on Matrix Analysis and Applications*, 20:149–158, 1998.
- [11] L. Elsner. An optimal bound for the spectral variation of two matrices. *Linear Algebra and its Applications*, 71:77–80, 1985.
- [12] M. A. Fiol. Eigenvalue interlacing and weight parameters of graphs. *Linear Algebra and its Applications*, 290:275–301, 1999.
- [13] M. Grohe, K. Kersting, M. Mladenov, and E. Selman. Dimension reduction via colour refinement. In *Algorithms - ESA 2014 - 22th Annual European Symposium, Wroclaw, Poland, September 8-10, 2014. Proceedings*, pages 505–516, 2014.
- [14] E. V. Haynsworth. Quasi-stochastic matrices. *Duke Mathematical Journal*, 22(1), 1955.
- [15] E. V. Haynsworth. Applications of a theorem on partitioned matrices. *J. Res. Bur. Stand.*, 63B(2), 1959.

- [16] A. S. Householder. Unitary triangularization of a nonsymmetric matrix. *Journal of the ACM*, 5(4):339–342, 1958.
- [17] K. Kate and B. Ravindran. Epsilon equitable partition; a positional analysis method. In *Proceedings of the 15th International Conference on Management of Data, 2009, Mysore, India*. Computer Society of India, 2009.
- [18] W. Kuich. Quasi-block-stochastic matrices. *Pacific Journal of Mathematics*, 27(2):353–364, 1968.
- [19] R. B. Lehoucq. The computation of elementary unitary matrices. *ACM Transaction on Mathematical Software*, 22(4):393–400, 1996.
- [20] J. Lerner. Role assignments. In U. Brandes and T. Erlebach, editors, *Network Analysis*, volume 3418 of *Lecture Notes in Computer Science*, pages 216–252. Springer, Berlin / Heidelberg, 2005.
- [21] X. Sun. On elementary unitary and ϕ -unitary transformations. Technical report, Duke University, 1995.